

$x_1, x_2, x_3, \dots$   $\xrightarrow{a}$  'get close' to some function?

$a_1, a_2, a_3, \dots \in \mathbb{R}$

① almost sure  $\xleftarrow{\text{def}} \text{SLLN}$

② conv. in prob.  $\xleftarrow{\text{def}} \text{WLLN}$

③ conv. in distri  $\xleftarrow{\text{def}} \text{CLT}$

CLT: iid rv  $X_1, X_2, \dots, X_n$ ,  $E(X_i) = \mu$ ,  $\text{Var}(X_i) = \sigma^2$

$S_n \sim N\left(E[S_n], \text{Var}[S_n]\right)$ , where  $S_n = \sum_{i=1}^n X_i$

$$M_n = \frac{S_n}{n}$$

↓

$$Z \sim \frac{S_n - E[S_n]}{\sqrt{\text{Var}[S_n]}} \sim N(0, 1)$$

$$\Pr(Z \leq z) = \Phi(z)$$

### 1. Confidence Interval Comparisons

In order to estimate the probability of a head in a coin flip,  $p$ , you flip a coin  $n$  times, where  $n$  is a positive integer, and count the number of heads,  $S_n$ . You use the estimator  $\hat{p} = \bar{S}_n/n$ .

(a) You choose the sample size  $n$  to have a guarantee

$$\Pr(|\hat{p} - p| \geq \delta) \leq \epsilon.$$

Using Chebyshev Inequality, determine  $n$  with the following parameters. Note that you should not have  $p$  in your final answer.

- (i) Compare the value of  $n$  when  $\epsilon = 0.05$ ,  $\delta = 0.1$  to the value of  $n$  when  $\epsilon = 0.1$ ,  $\delta = 0.1$ .
- (ii) Compare the value of  $n$  when  $\epsilon = 0.1$ ,  $\delta = 0.05$  to the value of  $n$  when  $\epsilon = 0.1$ ,  $\delta = 0.1$ .

(b) Now, we change the scenario slightly. You know that  $p \in (0.4, 0.6)$  and would now like to determine the smallest  $n$  such that

$$\Pr\left(\frac{|\hat{p} - p|}{p} \leq 0.05\right) \geq 0.95.$$

Use the CLT to find the value of  $n$  that you should use. Recall that the CLT states that the sum of IID random variables tends to a normal distribution with the sample mean and variance as its parameters for  $n$  large enough.

$$a) \Pr\left(\frac{|\bar{X} - E(X)|}{\sigma/\sqrt{n}} \geq \epsilon\right) \leq \frac{\text{Var}(X)}{\epsilon^2}$$

$$E(\bar{X}) = \frac{1}{n} E(S_n) = \frac{1}{n} \times np = p$$

$$\Rightarrow \Pr(|\hat{p} - p| \geq \epsilon) \leq \frac{\text{Var}(\hat{p})}{\epsilon^2}$$

$$\text{Var}(\hat{p}) = \frac{1}{n} \text{Var}(S_n) = \frac{1}{n} \times n \text{Var}(S) = \frac{1}{n} \sigma^2 = \frac{p(1-p)}{n}$$

$$\frac{d}{dp} p(1-p) = 0$$

$$1 - 2p = 0 \Rightarrow p = \frac{1}{2}$$

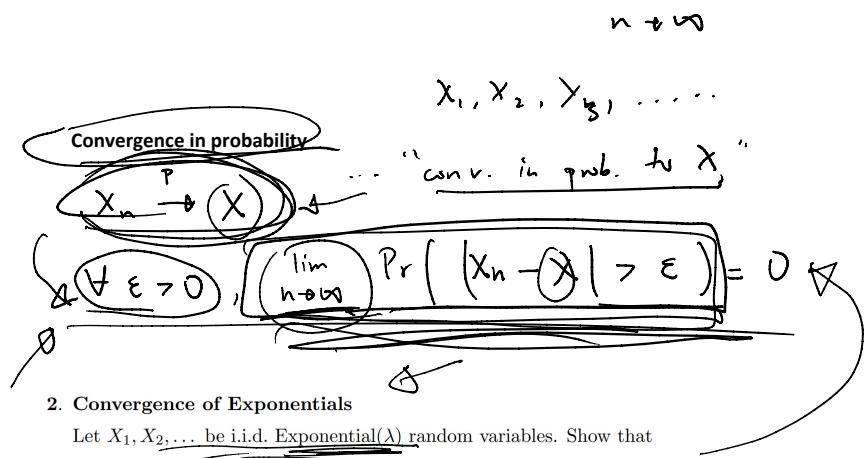
$$\frac{1}{n} \frac{p(1-p)}{\epsilon^2} \leq \delta \Rightarrow \frac{1}{4\epsilon^2} \leq n$$

$$S_i \sim \text{Bern}(p)$$

$$\text{Var}(S_i) = \sigma^2 = p(1-p)$$

$$\sigma^2 = p(1-p)$$

$$\text{Var}(\bar{S}_n) = \frac{1}{n^2} \text{Var}(S_n) = \frac{1}{n^2} \times n \text{Var}(S_1) = \frac{1}{n} \sigma^2 = \frac{\sigma^2(1-p)}{n}$$



## 2. Convergence of Exponentials

Let  $X_1, X_2, \dots$  be i.i.d.  $\text{Exponential}(\lambda)$  random variables. Show that

$X_n \xrightarrow{\ln n} 0$  in probability as  $n \rightarrow \infty$ .

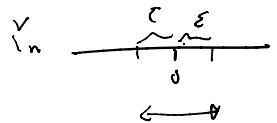
$X'_1 = \frac{X_1}{\ln 1}, X'_2 = \frac{X_2}{\ln 2}, \dots, X'_n = \frac{X_n}{\ln n}$

$\gamma_n = \frac{X_1}{\ln 1}, \gamma_2 = \frac{X_2}{\ln 2}, \dots, \gamma_n = \frac{X_n}{\ln n}$

Our sequence of RVs

is  $\gamma_1, \gamma_2, \gamma_3, \dots$ , where  $\gamma_i = \frac{X_i}{\ln i}$

$$f(w) = \frac{1}{2^n} \quad f(w) \rightarrow 0 \quad \lim_{n \rightarrow \infty}$$



$$\gamma_n \xrightarrow{P} 0$$

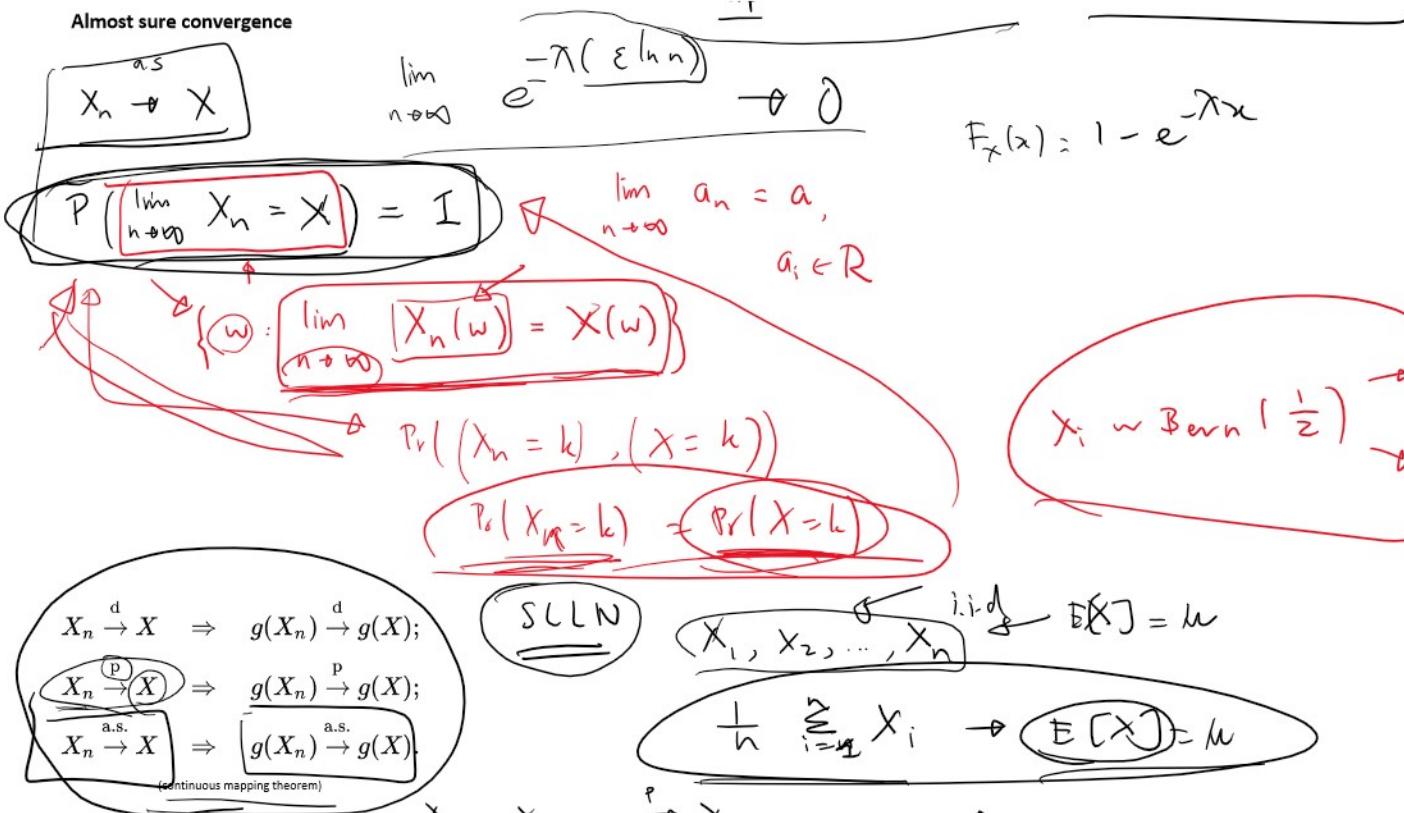
$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} \Pr(|\gamma_n - 0| > \epsilon) = 0$$

$$\begin{aligned} \Pr\left(\left|\frac{X_n}{\ln n}\right| - 0 > \epsilon\right) &= \Pr\left(\frac{X_n}{\ln n} > \epsilon\right) = \Pr(X_n > \epsilon \ln n) \\ &= 1 - \text{CDF}_{\text{exp}}\left(\frac{\epsilon \ln n}{\lambda}\right) = e^{-\lambda(\epsilon \ln n)} \end{aligned}$$

Almost sure convergence

$\lim_{n \rightarrow \infty} \gamma_n$

### Almost sure convergence



### 3. Breaking a Stick

I break a stick  $n$  times, where  $n$  is a positive integer, in the following manner: the  $i$ th time I break the stick, I keep a fraction  $X_i$  of the remaining stick where  $X_i$  is uniform on the interval  $[0, 1]$  and  $X_1, X_2, \dots, X_n$  are i.i.d. Let  $P_n = \prod_{i=1}^n X_i$  be the fraction of the original stick that I end up with.

- Show that  $P_n^{1/n}$  converges almost surely to some constant function.
- Compute  $E[P_n^{1/n}]$ .

$$\lim_{n \rightarrow \infty} P_n^{1/n} = \square$$

$$\lim_{n \rightarrow \infty} (\ln P_n^{1/n}) = \lim_{n \rightarrow \infty} \frac{1}{n} (\ln P_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \ln X_i$$

as  $\xrightarrow{\text{SLLN}}$   $E(\ln X_i)$

by SLLN

$$P_n^{1/n} \xrightarrow{\text{a.s.}} \square$$

$$(g(x) = e^x)$$

$$\lim_{n \rightarrow \infty} e^{\ln P_n^{1/n}} \xrightarrow{\text{a.s.}} e^{E(\ln X_i)}$$

$$\frac{\ln P_n^{1/n}}{E(\ln X_i)} \xrightarrow{\text{a.s.}} \int_{x=0}^1 \ln x dx$$

$$= \int_0^1 \frac{\ln x}{x} dx$$

$$= (x \ln x - x) \Big|_0^1 = 0 - 1 = -1$$

Markov's:  $x > 0, \Pr(X \geq c) \leq E[X]/c$

Chebyshov's:  $\Pr(|X - E(X)| \geq c) \leq \text{Var}[X]/c^2$

WLLN:  $X_1, X_2, \dots, X_n$  w/  $E[X_i] = \mu, \text{Var}[X_i] = \sigma^2$

define  $M_n \stackrel{\text{IID}}{\sim} \sum_{i=1}^n X_i, E[M_n] = n\mu, \text{Var}[M_n] = n\sigma^2$

$\Pr(|M_n - E[M_n]| \geq c) \leq \frac{\sigma^2}{c^2}$

n defines  $M_n = \frac{1}{n} \sum_{i=1}^n \lambda_i$ ,  $E[M_n] = \mu$ ,  $\text{Var}[M_n] = \frac{\sigma^2}{n}$

$$\Pr(|\text{empirical mean} - \text{true mean}| > c) \leq \frac{\sigma^2}{nc^2}$$

$$\lim_{n \rightarrow \infty} \Pr(|M_n - \mu| \geq c) \leq \frac{\sigma^2}{nc^2} \xrightarrow{n \rightarrow \infty} 0$$

$$= 6 - 7 = -1$$